Amendments to Claims

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims

1 (previously amended): A method of extracting classifying data from an audio signal, the method comprising the steps of:

(a) processing said audio signal into a perceptual representation of its constituent frequencies;

(b) processing said perceptual representation into at least one learning representation of said audio data stream;

(c) inputting at least one said learning representation into a multi-stage classifier, said multi-stage classifier comprising one or more first stage classifiers and a final stage metalearner classifier, the first stage classifiers receiving the learning representations and generating a metalearner vector which is utilized by the final stage metalearner classifier to generate the classification of said audio signal, whereby said multi-stage classifier extracts classifying data from said learning representations and outputs the classification of said audio signal.

2 (original): The method of extracting classifying data from an audio signal according to claim 1, wherein the step of processing the audio data into a perceptual representation of its constituent frequencies comprises calculating, for a time sample window of a digital representation of said audio signal, a Fast Fourier Transform function.

3 (original): The method of extracting classifying data from an audio signal according to claim 1, wherein the step of processing said perceptual representation into at least one learning representation further comprises dividing said perceptual representation into a plurality of time slices.

4 (currently amended): The method of extracting classifying data from an audio signal according to claim 3, wherein each of said time slices is ~~about~~ 0.8 to about 1.2 seconds in length.

5 (original): The method of extracting classifying data from an audio signal according to claim 1, wherein the step of dividing the perceptual representation into learning representations further comprises dividing said perceptual representation into a plurality of frequency bands.

6 (original): The method of extracting classifying data from an audio signal according to claim 5, wherein said plurality of frequency bands comprises 20 frequency bands.

7 (previously amended): The method of extracting classifying data from an audio signal according to claim 5, wherein the size of each of said frequency bands grows according to a golden ratio of frequency with respect to pitch.

8 (original): The method of extracting classifying data from an audio data stream according to claim 5, wherein no said frequency band includes any frequency greater than 11 kHz.

9 (previously amended): The method of extracting classifying data from an audio signal according to claim 1, wherein said first stage classifier of said multi-stage classifier comprises at least one Support Vector Machine.

10 (previously amended): The method of extracting classifying data from an audio signal according to claim 10, wherein said first stage classifier of said multi-stage classifier comprises at least one Support Vector Machine per category of classification.

11 (previously amended): The method of extracting classifying data from an audio signal according to claim 1, wherein said final stage metalearner classifier of said multi-stage classifier comprises a neural network.

12 (original): The method of extracting classifying data from an audio signal according to claim 11, wherein said neural network comprises at least one input node per category of classification, and further wherein said neural net comprises at least one output node per category of classification.

13 (original): The method of extracting classifying data from an audio signal according to claim 12, wherein said neural network comprises a hidden layer, wherein said hidden layer comprises at least as many nodes as the number of said input nodes.

14 (original): The method of extracting classifying data from an audio signal according to claim 11, wherein said neural network operates on a Gaussian activation function.

15 (currently amended): The method of extracting classifying data from an audio signal according to claim 1, wherein said classifying data comprises ~~at least one of~~ artist ~~and genre~~ data.

16 (original): The method of extracting classifying data from an audio signal according to claim 1, further comprising the step of converting said audio signal into a pulse code modulated digital bitstream.

17 (original): The method of extracting classifying data from an audio signal according to claim 1, further comprising the step of measuring the confidence of said classification by said multi-stage classifier.

18 (previously amended): A computer readable storage medium, storing therein a program of instructions for causing a computer to execute process of extracting classifying data from an audio signal, said process comprising the steps of:

(a) processing said audio signal into a perceptual representation of its constituent frequencies;

(b) processing said perceptual representation into at least one learning representation;

(c) inputting said learning representations of said audio data stream into a multi-stage classifier, said multi-stage classifier comprising one or more first stage classifiers and a final stage metalearner classifier, the first stage classifiers receiving the learning representations and generating a metalearner vector which is utilized by the final stage metalearner classifier to generate the classification of said audio signal, whereby said multi-stage classifier extracts classifying data from said learning representations and outputs the classification of said audio signal.

19 (withdrawn): A method of representing an audio signal for machine learning comprising:

(a) creating a perceptual representation of said audio signal by performing a frequency domain transform on at least one time-sampled window of a digital representation of said audio signal, said perceptual representation comprising component magnitudes of constituent frequency vectors that comprise said audio signal;

(b) calculating a magnitude of each constituent frequency vector within said audio signal;

(c) grouping each of said constituent frequency vectors into a number of frequency bands;

(d) calculating an average magnitude of said constituent frequency vectors within each of said frequency bands; and

(e) arranging said magnitudes into a learning representation.

20 (withdrawn): The method according to claim 19 wherein said frequency domain transform is a Fast Fourier Transform.

21 (withdrawn): The method according to claim 19 wherein an average magnitude of said constituent frequency vectors within each of said frequency bands further comprises an aggregate average magnitude over a plurality of said time-sampled windows.

22 (withdrawn): The method according to claim 21 where said plurality of time-sampled windows comprises 12 time-sampled windows.

23 (withdrawn): The method according to claim 19 wherein no said frequency band includes any frequency greater than 11 kHz.

24 (withdrawn): The method according to claim 19 wherein said frequency bands grow in size according to the golden ratio of frequency with respect to pitch.

25 (withdrawn): The method according to claim 19 further comprising the step of converting said audio signal into a pulse code modulated bitstream for processing by said frequency domain transform.

26 (withdrawn): A computer readable storage medium, storing therein a program of instructions for causing a computer to execute process of representing an audio signal for machine learning, said process comprising the steps of:

(a) creating a perceptual representation of said audio signal by performing a frequency domain transform on at least one time-sampled window of a digital representation of said audio signal, said perceptual representation comprising component magnitudes of constituent frequency vectors that comprise said audio signal;

(b) calculating a magnitude of each constituent frequency vector within said audio signal;

(c) grouping each of said constituent frequency vectors into a number of frequency bands;

(d) calculating an average magnitude of said constituent frequency vectors within each of said frequency bands; and

(e) arranging said magnitudes into a learning representation.

27 (previously amended): An apparatus for classifying an audio data stream comprising:

(a) a means for covering an audio data stream into a perceptual representation of its constituent frequencies;

(b) a means for dividing said perceptual representation into learning representations; and

(c) a multi-stage classifying means trained to distinguish among classifying categories of said audio data stream, wherein said multi-stage classifying means further comprises one or more first stage classifying means and a final stage metalearner classifying means, said first stage classifying means receiving the learning representations and generating a metalearner vector which is utilized by the final stage metalearner classifying means to generate the classification of said audio signal and outputs the classification of said audio signal.

28 (original): The apparatus according to claim 27, wherein the said means for covering an audio data stream into a perceptual representation of its constituent frequencies comprises means to perform a Fast Fourier Transform function on at least one time-sampled window digital representation of said audio stream.

29 (original): The apparatus according to claim 27, wherein a means for dividing said perceptual representation into learning representations further comprises means for dividing said perceptual representation into a plurality of time slices.

30 (currently amended): The apparatus according to claim 29, wherein each of said time slices is ~~about~~ 0.8 to about 1.2 seconds in length.

31 (original): The apparatus according to claim 27, wherein said means for dividing said perceptual representation into learning representations further comprises means for dividing said perceptual representation into a plurality of frequency bands.

32 (original): The apparatus according to claim 31, wherein said plurality of frequency bands comprises 20 frequency bands.

33 (previously amended): The apparatus according to claim 31, wherein the size of each of said frequency bands grows according to a golden ratio of frequency with respect to pitch.

34 (original): The apparatus according to claim 31, wherein no said frequency includes any frequency higher than 11 kHz.

35 (previously amended): The apparatus according to claim 27, wherein said first stage classifier of said multi-stage classifier comprises at least one Support Vector Machine.

36 (previously amended): The apparatus according to claim 36, wherein said first stage classifier of said multi-stage classifier comprises at least one Support Vector Machine per category of classification.

37 (previously amended): The apparatus according to claim 27, wherein said final stage metalearner classifier of said multi-stage classifier comprises a neural network.

38 (original): The apparatus according to claim 37, wherein said neural network comprises at least one input node per category of classification, and further wherein said neural net comprises at least one output node per category of classification.

39 (original): The apparatus according to claim 38, wherein said neural network comprises a hidden layer, wherein said hidden layer comprises at least as many nodes as the number of said input nodes.

40 (original): The apparatus according to claim 37, wherein said neural network operates on a Gaussian activation function.

41 (currently amended): The apparatus according to claim 27, wherein said classifying categories comprise ~~at least one of~~ artist ~~and genre~~ data.

42 (original): The apparatus according to claim 27, further comprising a means to convert said audio signal into a pulse code modulated digital bitstream.

43 (original): The apparatus according to claim 27, further comprising a means for measuring the confidence of said classification by said multi-stage classifier.

44 (withdrawn): An apparatus for representing an audio signal for machine learning comprising:

(a) a means for performing a frequency domain transform on at least one time-sampled window of a digital representation of said audio signal, said perceptual representation comprising component magnitudes of constituent frequency vectors that comprise said audio signal;

(b) a means for calculating a magnitude of each constituent frequency vector;

(c) a means for grouping each of said constituent frequency vectors into a number of frequency bands;

(d) a means for calculating an average magnitude of said constituent frequency vectors within each of said frequency bands; and

(e) a means for arranging said magnitudes into a learning representation.

45 (withdrawn): The apparatus according to claim 44 wherein said means for performing a frequency domain transform comprises a means for performing a Fast Fourier Transform.

46 (withdrawn): The apparatus according to claim 44 wherein no said frequency band includes any frequency greater than 11 kHz.

47 (withdrawn): The apparatus according to claim 44 wherein said frequency bands grow in size according to the golden ratio of frequency with respect to pitch.

48 (withdrawn): The apparatus according to claim 44 further comprising a means for converting said audio signal into a pulse code modulated bitstream for processing by said frequency domain transform.